

An aerial photograph of the University of Exeter campus, showing green lawns, trees, and buildings, partially framed by a blue curved graphic element.

Economics Department Discussion Papers Series

ISSN 1473 – 3307

Updating Choquet Beliefs

Jürgen Eichberger, Simon Grant and David Kelsey

Paper number 06/07

September 20, 2005

Updating Choquet Beliefs.*

Abstract

We apply Pires's coherence property between unconditional and conditional preferences that admit a CEU representation. In conjunction with consequentialism (only those outcomes on states which are still possible can matter for conditional preference) this implies that the conditional preference may be obtained from the unconditional preference by taking the Full Bayesian Update of the capacity.

Attitudes towards sequential versus simultaneous resolution of uncertainty for a simple bet are analyzed. We show that for a class of recursive CEU preferences which exhibit both optimism and pessimism, a 'good-news' signal is preferred to no signal which is preferred to a 'bad-news' signal.

Keywords: updating ambiguous beliefs, Full Bayesian Updating, Choquet Expected Utility, optimism, pessimism, recursive preferences.

JEL classifications: D81

Jürgen Eichberger
Universität Heidelberg

Simon Grant
Rice University

David Kelsey
Exeter University

*We thank participants at the RUD 2005 meetings and in particular, Paulo Ghirardato, Eran Hanany, Robert Kast, Peter Klibanoff, Mark Machina and Marciano Siniscalchi for their comments and suggestions. This research is in part supported by ESRC grant RES-000-22-0650.

1 Introduction

In recent years there has been a mounting challenge to the standard model of decision under uncertainty, the subjective expected utility model. In particular, models in which beliefs cannot be represented by a single probability measure over the set of events, have been introduced as a way to formalize the distinction drawn by Knight (1921), Keynes (1921) and others between situations of risk (where probabilities are based upon an extensive data base of past relevant cases or can be readily gleaned from the structure and nature of the situation) and uncertainty (where probabilities are not well-known or agreed upon). For example, in the Choquet Expected Utility model (Schmeidler, 1989) beliefs are represented by a capacity, a measure that is not necessarily additive, while in the multiple-prior model (Gilboa & Schmeidler, 1989), as its name suggests, beliefs are represented by a set of probability measures.

There has been some success in applying these static models to individual decision making problems as well as many-agent settings including those involving strategic interaction (see for example, Dow & Werlang, 1992, Mukerji, 1998, Mukerji & Tallon, 2001, 2004 and Eichberger & Kelsey, 2000, 2002). But in order to be able to apply non-additive beliefs models to sequential or dynamic settings requires a theory of how preferences are updated as new information arrives.

In subjective expected utility, the almost universally applied rule is Bayesian updating. For non-additive beliefs there have been two major approaches in the literature. The first is a statistical approach that considers for different updating rules the statistical properties of the updated beliefs that are derived from such rules. Examples include Denneberg (1994, 2002), Jaffray (1992), Lapied & Kast (2005), Lehrer (2003) and Shafer (1976). The other approach is decision-theoretic. The updating rule arises from axioms on the preferences both unconditional and conditional: a non-exhaustive list of examples includes Epstein & Schneider (2001), Gilboa & Schmeidler (1993), Klibanoff & Hanany (2004), Pires (2002), Siniscalchi (2004), Sarin & Wakker (1998), Walley (1991) and Wang (2003). This paper

follows the decision-theoretic approach but we shall show in the sequel how the rule we obtain is the Full Bayesian Updating rule of Jaffray (1992) and coincides with the conditional expectation derived by Denneberg (2002).

For purposes of setting the benchmark against which alternatives will be introduced, motivated and developed, let us first consider how the Bayesian updating rules for subjective expected utility preferences arises from the following natural method to deduce the conditional preferences from the prior preferences. Suppose we wish to deduce from the prior preference relation the preference between two acts, say f and g , conditional on the event E having obtained. Savage invoking his ‘sure-thing principle’ would argue it is enough to look at the unconditional preference between any pair of acts, say f' and g' which *agree* on states outside the conditioning event E , whereas for states in E , f' agrees with f and g' agrees with g . Since f' and g' agree on the states outside the conditioning event, Savage argues it is reasonable to assume that only how they differ on states in E will (or should) be decisive in determining the (unconditional) preference between them. If f' is preferred to g' unconditionally, then conditional on knowing that E has obtained, the individual should also prefer f to g (and indeed she should also still prefer f' to g'). For the unconditional preference, as f' and g' agree on states outside E , should E not obtain, then both acts will lead to the same outcome. For the conditional preference relation, knowing that E has obtained makes what f and g might have led to on states outside of E , immaterial.

Such reasoning embodies two properties:

1. consequentialism – only those outcomes on states which are still possible can matter for preference; and
2. dynamic consistency – if for an unconditional preference relation we have one act is preferred to another, then conditional on knowing that the complement of an event on which the two acts agree has obtained, the conditional preference should also have the latter act preferred to the former.

So in our example, consequentialism requires that if conditional on knowing that E has

obtained, f is preferred to g , then the conditional preference relation should also have f' be preferred to g' . Dynamic consistency requires that if the unconditional preference relation has f' preferred to g' , then the conditional preference relation of the individual after she learns that E has obtained, should also have f' preferred to g' .

In the extant literature, Hanany & Klibanoff (2005), Siniscalchi (2004) and Sarin & Wakker (1998) drop consequentialism and retain dynamic consistency. Epstein & Schneider (2001) show it is possible to retain both if one restricts the domain of acts and conditioning events (or more precisely, filtrations) over which preferences are defined. Wang (2003) casts his analysis in a more complicated setting of consumption–information profiles which have no direct counter-part in a standard Savage act framework, but effectively he is imposing similar restrictions to those of Epstein & Schneider on the domain of admissible problems.¹ We shall, however, maintain an unrestricted domain of acts and conditioning events and follow Gilboa & Schmeidler (2003), Pires (2002) and Walley (1991) in retaining consequentialism and dropping dynamic consistency.

Our reason for retaining consequentialism and dropping dynamic consistency is because, to paraphrase Baron & Frisch (1988), we feel ambiguity arises in a fundamental sense from uncertainty about probability created by missing information that is relevant and could be known. Hence once an event is known to have obtained, the only remaining ambiguity the individual faces relates to uncertainty about the probabilities of subevents of that event. Past (or borne) uncertainty one may have had about the probability of counterfactual event and its subsets are no longer relevant. But such uncertainty might have been relevant to the individual at the time when she did not know whether the event or its complement had obtained, and so such ambiguity that she perceived there to have been *ex ante*, may well have had an impact on her unconditional preferences. As we shall see now this may well lead to violations of dynamic consistency.

¹ Eichberger et al (2005) also restrict preferences over information structure for a fixed filtration. But for the particular family of non-additive measures they consider for beliefs, they show a necessary and sufficient condition for dynamic consistency is that beliefs be additive over the final stage in the filtration.

As an illustration, consider an Ellsberg type urn that contains one hundred balls numbered 1 to 200. Suppose the balls numbered 1 to 66 are red. The balls numbered from 67 to $200 - 2n$ are black and the remainder (that is, those numbered from $201 - 2n$ to 200) are white. The only information the decision maker has about n is that it is an integer no smaller than one and no larger than sixty-six. That is, she knows the number of black balls is an even number but it could be as few as two or as many as one hundred and thirty-two, similarly, for the number of white balls. A ball is to be drawn randomly from the urn, and the decision maker has to choose among different bets concerning the color and number of the ball drawn. Suppose there are two possible outcomes, ‘win’ or ‘lose’, thus any act may be characterized by the event on which a win will arise.

First consider her preference between the act ‘win if (and only if) the ball drawn is red’ and a second act ‘win if the ball drawn is black’. Given her information, the individual knows that she will win with the first act if any one of the sixty-six red balls in the urn is drawn, but with the second act she only knows that there could be any even number of black balls from two to one hundred and thirty-two. If she is averse to bets with ambiguous odds, we may well expect her to express an unconditional strict preference for the first act over the second.

Now consider her preference between the first act ‘win if the ball drawn is red’, and a third act ‘win if either the ball drawn has an odd number and is black or it has an even number and is white’. Given her information, she knows that she will win with the third act if any one of the sixty-seven balls that are odd and black or that are even and white is drawn. Thus we would expect her to express an unconditional strict preference for the third act over the first. And if her preferences are transitive, then we would expect her to choose the third act if the set of options available to her comprised just those three acts.

So suppose she chooses the third act, a ball is then drawn out of the urn and she is told its number is odd. But before its color is announced she is told she may switch her bet to either of the other two acts if she so wishes. Given the information that the ball drawn has an odd number on it, she knows that out of the one hundred balls with odd numbers,

thirty-three are red, but out of the remaining sixty-seven balls all she knows is that at least one of them is black and at least one of them is white. So what might her conditional preferences be? We suggest it is reasonable to assume that given she knows the ball drawn has an odd number on it, that she will be indifferent between the second and third acts, since both yield a win if and only if the odd numbered ball that has been drawn is black. Furthermore, we suggest that since her expressed unconditional strict preference for the first act over the second act, revealed an aversion to bets with ambiguous odds, we would expect her, conditional on knowing that the number of the ball drawn is odd, to prefer the first act over the third act. If these were her conditional preferences then she would switch from the third act to the first act (that is, switch to a bet on red rather than retain her bet on black). Such behavior violates the notion of dynamically consistency defined above, but we feel it is more in accord with the spirit of the original Ellsberg paradox.

Our principal aim in this paper is to axiomatize an updating rule for a Choquet Expected Utility maximizer (that is, the extension of subjective expected utility where beliefs are represented by a capacity rather than an [additive] probability measure). After establishing the analytical framework in Section 2, we introduce, in Section 3, the axiom Pires (2002) proposed to link the unconditional and conditional preferences. This axiom, which we dub, *Conditional Certainty Equivalent Consistency*, has a similar intuitive appeal to Savage's sure-thing principle but is weak enough to accommodate standard Ellsberg type behavior, including that described in the example above. The representation result that we derive for the case where the unconditional and conditional preferences are all members of the Choquet Expected Utility family of preferences, shows that the utility function over outcomes is invariant to updating, while the updated capacity may be obtained using Jaffray's (1992) Full Bayesian updating rule (or Walley's, 1991, Generalized Bayesian updating rule) on the unconditional capacity. We then show in Section 5 that this rule also coincides with Denneberg's (2002) proposed definition of conditional expectation for Choquet integration. Comparisons are made in section 6 between this rule and two other rules that have been axiomatized for Choquet Expected utility preferences, one of which is the much celebrated

Dempster-Shafer updating rule. Our main insight is that only the Full Bayesian updating rule preserves mixed optimistic and pessimistic attitudes towards ambiguity that are present in the unconditional preference relation. For the other two rules the updated capacity exhibits either extreme optimism or (as is the case with the Dempster-Shafer rule) extreme pessimism, regardless of what the attitude was in the unconditional preferences.

We conclude with an application analysing attitudes towards the sequential versus the simultaneous resolution of uncertainty for a simple bet. We show that for a particular class of Choquet expected utility maximizers who overweight the events that yield both the best and worst outcomes, the sequential resolution of uncertainty is more attractive for bets with gains that the individual perceives there to be a low likelihood of occurring and conversely, all-in-one resolution of uncertainty is more attractive for bets that she perceives there to be a low likelihood of loss (and hence high likelihood of no loss). Intriguingly this seems to correspond to ‘hope’ for low likelihood good events making sequential resolution of uncertainty more desirable, as well as ‘dread’ of low likelihood bad events making simultaneous resolution of uncertainty more desirable. Indeed for this particular class of Choquet expected utility maximizers, we show formally that ‘good-news’ signals (that is, signals in which receipt of the favorable signal realization means the bet definitely pays off) are strictly preferred to having no signal which in turn is strictly preferred to ‘bad-news’ signals (that is, signals in which receipt of the less favorable signal realization means the bet definitely does not pay off).

2 Setup

We present our analysis in the context of a framework of purely subjective uncertainty. We take the uncertainty a decision maker faces to be described by a finite set of *states*, denoted by S . Associated with the set of states is the set of events, taken to be the set of subsets of S , denoted by \mathcal{E} . For each $E \in \mathcal{E}$, E^c shall denote its complement.

Let X , the set of outcomes, be a connected and separable topological space. An *act* is

a function $f : S \rightarrow X$. \mathcal{F} denotes the set of such acts and is endowed with the product topology induced by the topology on X . We shall identify each $x \in X$ with the constant act, $f(s) = x$ for all $s \in S$. For any pair of acts f, g in \mathcal{F} and any event $E \in \mathcal{E}$, $f_E g$ will denote the act $h \in \mathcal{F}$, formed by ‘splicing’ the two acts f and g , in which $h(s)$ equals $f(s)$ if $s \in E$, and equals $g(s)$ if $s \notin E$. In general, for any finite partition $\{E_1, \dots, E_n\}$ of S and any list of n acts (f^1, \dots, f^n) , let $f_{E_1}^1 f_{E_2}^2 \dots f_{E_{n-1}}^{n-1} f^n$ be the act that yields $f^i(s)$ if s is in E_i .

We assume that the decision maker is characterized by a family of conditional preference relations on \mathcal{F} . For each event $E \in \mathcal{E}$, let \succsim_E denote the preferences over acts *given* E . That is, we shall interpret \succsim_E as the agent’s preferences if she knew that E had obtained. As usual \succ_E and \sim_E will denote the asymmetric and symmetric parts of \succsim_E , respectively. The relation \succsim shall denote the individual’s unconditional preference relation on \mathcal{F} (that is, $\succsim = \succsim_S$).

We say f and g are comonotonic if for every pair of states s and s' in S , $f(s) \succ f(s')$ implies $g(s) \succ g(s')$. Given a preference relation \succsim_E , an event $A \in \mathcal{E}$ is \succsim_E -null if $f_A g \sim_E g$ for all pairs of acts $f, g \in \mathcal{F}$. Let \mathcal{N}_E denote the set of \succsim_E -null events (and \mathcal{N} denote the set of [unconditional] null events, that is, $\mathcal{N} = \mathcal{N}_S$).

For ease of exposition (and without any essential loss of generality) we shall assume the existence of a best and worst outcome, namely that there exist outcomes 0 and M in X , such that $M \succ 0$ and $M \succsim x \succsim 0$ for all $x \in X$.

3 Connecting Conditional and Unconditional Preferences

The first property we require for the conditional preferences is *consequentialism*. That is, the conditional preferences are ‘forward-looking’ in the sense that what happens off the conditioning event should not be able to affect the conditional preference between any pair of acts.

Axiom 1 (*Consequentialism*) Fix an event $E \in \mathcal{E}$. The event E^c is \succsim_E -null. That is, $f_E g \sim_E f$ for all $f, g \in \mathcal{F}$.

This is a particularly desirable property to have in order to keep the preference model tractable in applications, since it means for a given conditional preference relation \succsim_E , we do not need explicitly to keep track of what outcomes might have resulted from an act had a state outside of E obtained.²

The next two axioms connect the conditional to the unconditional preferences. The first simply requires that the ordering of outcomes be the same across states. Notice that in conjunction with the existence of an unconditionally best outcome and worst outcome, this entails that every \succsim_E is non-degenerate.

Axiom 2 (*State Independence*) For any pair of outcomes x, y in X , and any event E in \mathcal{E} , $x \succ y$ if and only if $x \succsim_E y$.

Our third axiom is taken from Pires (2002). It says that if conditional on E obtaining, the decision maker is indifferent between the act f and the outcome x , then her unconditional preferences should also express indifference between the outcome x and the act $f_E x$, that is the act that agrees with f on E and agrees with x on the complement of E .

² The fact that by definition the conditional preference relation \succsim_E does not depend on the act the individual may have chosen ex ante, already embodies much of the force of consequentialism. More properly, consequentialism should be viewed as the joint hypothesis that \succsim_E does not depend on the act the individual chose ex ante and that the event E^c is \succsim_E -null. For example, Hanany & Klibanoff (2005) relax consequentialism by *dropping* the former while retaining the latter.

Axiom 3 (Conditional Certainty Equivalent Consistency) For any unconditionally non-null event $E \notin \mathcal{N}$ any outcome x in X , and any act f in \mathcal{F} , $f \sim_E x$ implies $f_E x \sim x$.

It is reminiscent of Savage’s sure-thing principle. If on knowing that E had obtained, an individual would be indifferent between the act f and the outcome x , then from the perspective of his unconditional preferences he should be indifferent between x for sure, and the act $f_E x$, that is, the act that coincides with f on E and yields x should a state outside of E obtain. From the ex ante perspective, one might justify this with the following reasoning: suppose the decision maker runs the thought experiment in which he imagines he is going to learn whether the state of the world is in E or is not. If he learns the state is in E , then he knows he will be indifferent between $f_E x$ and x . On the other hand if he learns it is not in E , then he knows he will receive x for sure. Hence, since he anticipates he will be either in a situation in which he is indifferent between $f_E x$ and the outcome x or in a situation in which he receives x for sure, he reasons he should be indifferent between $f_E x$ and x now, when he does not know whether the state is in E or its complement.

A stronger requirement is the following property that appears in Skiadas (1997).

Strict Coherence For any non-null event $E \notin \mathcal{N}$ and any pair of acts g and h in \mathcal{F} :

1. $g \succsim_E h$ and $g \succsim_{E^c} h$ implies $g \succsim h$.
2. $g \succ_E h$ and $g \succeq_{E^c} h$ implies $g \succ h$.

Strict coherence precludes any ‘hedging’ benefits (or costs) that might be associated with the unconditional preferences. Indeed, in conjunction with consequentialism, strict coherence essentially entails that the unconditional preferences are additively separable across states (see Skiadas, 1997, Section 4 for details). Conditional Certainty Equivalent Consistency does not rule out non-neutral attitudes towards ‘hedging’ because when we consider the implication for the unconditional preference, we do so only for acts that are constant and agree on the complementary event. Roughly speaking, Conditional Certainty Equivalent

Consistency is the restriction of Strict Coherence to pairs of acts g and h , where $h = x$ is a constant act, and $g(s) = x$ for all $s \notin E$.

To see why strict coherence might not hold in the context of ambiguous beliefs, recall the Ellsberg type urn example discussed in the introduction. The urn contains one hundred balls numbered 1 to 200. The balls numbered 1 to 66 are red. The balls numbered from 67 to $200 - 2n$ are black and the remainder (that is, those numbered from $201 - 2n$ to 200) are white. The only information the decision maker has about n is that it is an integer and that $1 \leq n \leq 66$.

Let O (respectively, E) be the event that the ball drawn from the urn has an odd (respectively, even) number on it. Let R (respectively, B , W) be the event that the ball drawn is red (respectively, black, white) in color. Let OR be the event that the ball drawn from the urn has an odd number and its color is red, and so on. Recall the pair of acts, discussed in the introduction: $g = M_R 0$ and $h = M_{OB} M_{EW} 0$. Given her information, the individual knows that she will win with g if any one of the sixty-six red balls in the urn is drawn. Similarly, she knows that she will win with h , if any one of the sixty-seven balls that are odd and black or that are even and white is drawn. Thus we would expect for her unconditional preferences she would have $h \succ g$.

But what about her conditional preferences relations \succsim_O and \succsim_E ? If she knows that the ball drawn has an odd number on it, then out of the one hundred balls with odd numbers, she knows thirty-three are red, but out of the remaining sixty-seven balls all she knows is that at least one of them is black and at least one of them is white. If she is averse to bets with ambiguous odds, we may well expect for conditional preference relation to have $g \succ_O h$, and by similar reasoning $g \succ_E h$. A violation of Strict Coherence.

On the other hand, if she expressed conditional on knowing the number of the ball drawn was odd, an indifference between the act $f = M_B 0$ (that is, betting on the color of the ball drawn being black) and say the amount for sure $x = 0.30 \times M$, that is, $f \sim_O x$, then Conditional Certainty Equivalent Consistency requires her to express an unconditional indifference between the act $f_O x = M_{OB} 0_{OR} 0_{OW} x$ and getting x for sure, that is $f_O x \sim x$.

But this by itself does not seem to contradict any of our intuitions for the individual's conditional or unconditional preferences.

4 The Representation Result

Pires conducted her analysis in the context of the multiple priors model of Gilboa & Schmeidler (1989). That is, she imposed the axioms of Gilboa & Schmeidler (1989) so that each preference relation \succsim_E admitted a representation of the form

$$f \succsim_E g \Leftrightarrow \min_{p \in \Delta_E} \int u \circ f \, dp \geq \min_{p \in \Delta_E} \int u \circ g \, dp,$$

where Δ_E is a convex set of probability measures over S , with the property: $p \in \Delta_E$ implies $p(E) = 1$. Let $\Delta = \Delta_S$, that is, the set of priors for the unconditional preference relation \succsim . Pires shows that if such a family of multiple prior preferences satisfy *Conditional Certainty Equivalent Consistency* then for any event E and any $p \in \Delta$, such that $p(E) > 0$, there exists $q \in \Delta_E$, where q is the Bayesian update of p , that is,

$$q(A) = \frac{p(A \cap E)}{p(E)}, \text{ for any } A \in \mathcal{E}.$$

That is, for any prior probability that gives positive weight to the conditioning event, its Bayesian update is an element of the updated set of multiple priors. Furthermore, if for every $p' \in \Delta$, $p'(E) > 0$, then for every $q \in \Delta_E$, there exists $p \in \Delta$, such that q is the Bayesian update of p . That is, if every prior probability gives positive weight to the conditioning event, then the set of updated priors is obtained by updating all the prior probability measures using Bayes' rule.

In this section we explore what the axioms introduced in the previous section imply for a family of preferences in which each preference relation \succsim_E admits a Choquet Expected Utility (CEU) representation. The Choquet Expected Utility of an act f , is taken with respect to a utility index over outcomes, $u : X \rightarrow \mathbb{R}$, and a normalized and monotonic set function (or capacity) $\nu : \mathcal{E} \rightarrow [0, 1]$, that satisfies $\nu(\emptyset) = 0$, $\nu(S) = 1$ and $A \subset B \Rightarrow \nu(A) \leq \nu(B)$, for all $A, B \in \mathcal{E}$.

Definition 1 Fix a capacity $\nu : \mathcal{E} \rightarrow [0, 1]$. The conjugate capacity, denoted $\bar{\nu}$, is defined as $\bar{\nu}(E) = 1 - \nu(E^c)$.

Since every act is finite-ranged, for each act f , we can find a finite partition $\{E_1^f, \dots, E_n^f\}$ of S and a list of n outcomes (x_1^f, \dots, x_n^f) , such that $u(x_i^f) \geq u(x_{i+1}^f)$, $i = 1, \dots, n-1$ and $f = (x_1^f)_{E_1^f} (x_2^f)_{E_2^f} \dots (x_{n-1}^f)_{E_{n-1}^f} (x_n^f)$. Formally, the Choquet Expected Utility of an act f with respect to the utility index u and the capacity ν may be defined as:

$$\int u \circ f \, d\nu = \nu(E_1^f) u(x_1^f) + \sum_{i=2}^n \left[\nu\left(\bigcup_{j=1}^i E_j^f\right) - \nu\left(\bigcup_{j=1}^{i-1} E_j^f\right) \right] u(x_i^f). \quad (1)$$

In particular, the Choquet expected utility of the simple bet $x_E y$, where $x \succsim y$, is given by

$$\int u \circ (x_E y) \, d\nu = \nu(E) u(x) + \bar{\nu}(E^c) u(y).$$

Definition 2 (CEU Preferences) The set of conditional preference relations $\langle \succsim_E \rangle_{E \in \mathcal{E}}$ is said to constitute a collection of CEU preferences, if for each \succsim_E , there exists a capacity ν_E on \mathcal{E} and a continuous non-constant real-valued function u_E on X such that for all $f, g \in \mathcal{F}$

$$f \succsim_E g \Leftrightarrow \int u_E \circ f \, d\nu_E \geq \int u_E \circ g \, d\nu_E.$$

As is well-known, a CEU preference relation admits a multiple prior representation if (and only if) the capacity is convex, that is, for all pairs of events A and B ,

$$\nu(A \cup B) \geq \nu(A) + \nu(B) - \nu(A \cap B).$$

Thus the intersection of the two models is non-empty, but not all multiple prior preferences admit a CEU representation and clearly, since not all capacities are convex, not all CEU preferences admit a multiple prior representation. So neither family is a special case of the other. Furthermore, non-convex capacities can be used to model individuals who have both optimistic as well as pessimistic attitudes towards ambiguity (see Wakker, 2001).

We can now state and prove our main representation result: a family of CEU preferences satisfy the three axioms above if and only if the all the utility indices are the same

(up to normalization) and the capacity for the conditional preference is obtained from the unconditional capacity by the Full Bayesian Updating (FBU) rule of Jaffray (1992).³

Theorem 1 *Fix $\langle \succsim_E \rangle_{E \in \mathcal{E}}$ a collection of CEU preferences. For each $E \in \mathcal{E}$, let (u_E, ν_E) be the utility index and neo-additive capacity associated with \succsim_E . The following two statements are equivalent:*

1. $\langle \succsim_E \rangle_{E \in (\mathcal{E})}$ satisfies Consequentialism, State Independence and Conditional Certainty Equivalent Consistency.
2. For each pair of events $E, B \in \mathcal{E}$, for which $\nu(B \cap E) > 0$ and $\bar{\nu}(B^c \cap E) > 0$ there exists some $\alpha_E > 0$, $\beta_E \in \mathbb{R}$, such that

$$\begin{aligned}
 u_E(x) &\equiv \alpha_E u(x) + \beta_E, \\
 \text{and for all } A \in \mathcal{E}, \nu_E(A) &\equiv \frac{\nu(A \cap E)}{\nu(A \cap E) + \bar{\nu}(A^c \cap E)}, \\
 \text{whenever} &\quad \nu(A \cap E) + \bar{\nu}(A^c \cap E) > 0.
 \end{aligned}$$

Remark 1 The theorem clearly extends Pires's updating result to the class of CEU preferences. But notice that we have done this in a setting of purely subjective uncertainty. Hence our state independence assumption is much weaker than hers, as it only entails that the ordinal ranking of outcomes remains unchanged no matter on which event preferences are being conditioned. In the Anscombe-Aumann setting of mixed subjective and objective uncertainty employed by Pires (2002), state independence says that the *certainty equivalent* of any pure roulette lottery is the same no matter on which state it obtains and no matter what obtains on other states. This immediately gives her the invariance of the 'risk preferences over pure roulette lotteries' and hence that the conditional utility functions are the same up to a positive affine transformation. The novel aspect of our proof is that provided not all subsets of the conditioning event are null or conditionally universal we are able to show the 'cardinal invariance' of the state utility functions over outcomes arises from Conditional

³ This is also referred to as the Generalized Updating Rule by Walley (1991).

Certainty Equivalent Consistency.⁴ That is, it is implied by the updating rule embodied in this axiom.

5 The Max-Min Representation of the Choquet Integral and Conditional Expectation

Denneberg (2002) shows that any capacity admits what he dubs a max-min representation. To formally state this representation we need to introduce the following concepts and definitions. A capacity μ is called k -**monotone**, $k \geq 2$, if for $A_1, \dots, A_k \subset S$

$$\mu\left(\bigcup_{i=1}^k A_i\right) + \sum_{I \subset \{1, \dots, k\}} (-1)^{|I|} \mu\left(\bigcap_{i \in I} A_i\right) \geq 0.$$

Convexity or 2-**monotonicity** corresponds to: for any pair of events A and B ,

$$\mu(A \cup B) + \mu(A \cap B) \geq \mu(A) + \mu(B).$$

A capacity is *totally monotone* or a *belief function* if it is k -**monotone** for any $k \geq 2$.

Fix a capacity ν . Define the *totally monotone core* of ν as

$$C_{\ll}(\nu) := \{\beta \mid \beta \text{ belief function on } 2^S, \beta \leq \nu\}$$

And for an arbitrary belief function β , define its (*additive*) *core* to be

$$C_+(\beta) := \{\alpha \mid \alpha \text{ additive on } 2^S, \alpha \geq \beta\}.$$

Denneberg (2002) shows that there exist *max-min additive representations* of ν and of the Choquet integral wrt ν :

$$\nu = \max_{\beta \in C_{\ll}(\nu)} \min_{\alpha \in C_+(\beta)} \alpha \text{ and } \int X d\nu = \max_{\beta \in C_{\ll}(\nu)} \min_{\alpha \in C_+(\beta)} \int X d\alpha.$$

His approach is then to start with conditional expectation for additive measures and generalize it by means of the max-min representation.

⁴ The requirement that not all subsets of the conditioning event are null or conditionally universal follows from the hypothesis that there exists an event B , for which $\nu(B \cap E) > 0$ and $\bar{\nu}(B^c \cap E) > 0$.

Let $E_\alpha[X|E]$ denote the conditional expectation of the real-valued function X with respect to the additive measure α given the event E has obtained. If $\alpha(E) > 0$ then $E_\alpha[X|E]$ is well-defined and equal to

$$E_\alpha[X|E] = \frac{1}{\alpha(E)} \int_E X d\alpha.$$

For the sequel, add the element ∞ to the real line which we shall take to be neutral with respect to the minimum and maximum simultaneously. That is, for all x in $\mathbb{R} \cup \{\infty\}$, $\min\{x, \infty\} = \max\{x, \infty\} = x$, $x + \infty = x - \infty = x \times \infty = \infty$. We shall set $E_\alpha[X|E] := \infty$, whenever $\alpha(E) = 0$.

Denneberg then defines the *conditional expectation* of X with respect to the capacity ν given the event E has obtained as

$$E_\nu[X|E] = \max_{\beta \in C_{\leq}(\nu)} \min_{\alpha \in C_+(\beta)} E_\alpha[X|E].$$

We can show that given an event E has obtained, for certain cases the conditional expectation of the indicator function of the event A is simply the updated capacity of the event A obtained using the Full Bayesian Updating Rule.

Proposition 1 *Fix a capacity ν and an event E for which $\nu(E) > 0$. The conditional expectation of the indicator function $\mathbf{1}_A$, $A \subset S$, wrt the capacity ν given E , is given by*

$$E_\nu[\mathbf{1}_A|E] = \frac{\nu(A \cap E)}{\nu(A \cap E) + \bar{\nu}(A^c \cap E)}, \text{ whenever } \nu(A \cap E) + \bar{\nu}(A^c \cap E) > 0.$$

Proof. Denneberg (2002, Example 4.3) shows that the result holds for the case where the capacity β is a belief function. That is, if $\beta(E) > 0$ we have

$$\begin{aligned} E_\beta[\mathbf{1}_A|E] &= \min_{\alpha \in C_+(\beta)} E_\alpha[X|E] \\ &= \min_{\alpha \in C_+(\beta)} \frac{\alpha(A \cap E)}{\alpha(E)} \\ &= \frac{\beta(A \cap E)}{\beta(A \cap E) + \bar{\beta}(A^c \cap E)}. \end{aligned}$$

The first equation follows from the min representation of the Choquet integral with respect to a belief function. The second equation is the standard Bayesian update of a probability,

the assumption $\beta(E) > 0$, guarantees the denominator does not vanish. The third equation follows from Denneberg (1994), Theorem 2.4.

So now let ν be a capacity (not necessarily a belief function) for which $\nu(E) > 0$ and $\nu(A \cap E) + \bar{\nu}(A^c \cap E) > 0$. It follows from the max-min representation of the Choquet integral and the first part of the proof, that,

$$\begin{aligned} E_\nu[\mathbf{1}_A|E] &= \max_{\beta \in C_{\ll}(\nu)} E_\beta[\mathbf{1}_A|E] \\ &= \max_{\beta \in C_{\ll}(\nu)} \frac{\beta(A \cap E)}{\beta(A \cap E) + \bar{\beta}(A^c \cap E)} \end{aligned}$$

The assumption $\nu(A \cap E) + \bar{\nu}(A^c \cap E) > 0$ implies either (i) $\bar{\nu}(A^c \cap E) > 0$ or (ii) if $\bar{\nu}(A^c \cap E) = 0$, then $\nu(A \cap E) > 0$. Clearly, if case (ii) holds we have by definition

$$\nu(A \cup E^c) = \max_{\beta \in C_{\ll}(\nu)} \beta(A \cup E^c) = 1$$

and so we have

$$E_\nu[\mathbf{1}_A|E] = 1 = \max_{\beta \in C_{\ll}(\nu)} \frac{\beta(A \cap E)}{\beta(A \cap E) + \bar{\beta}(A^c \cap E)}.$$

So now suppose case (i) holds. That is, $\beta(A \cup E^c) < 1$ or equivalently, $\bar{\beta}(A^c \cap E) > 0$ for all $\beta \in C_{\ll}(\nu)$. Hence, for any two belief functions $\beta, \beta' \in C_{\ll}(\nu)$

$$\begin{aligned} &\frac{\beta(A \cap E)}{\beta(A \cap E) + \bar{\beta}(A^c \cap E)} \geq \frac{\beta'(A \cap E)}{\beta'(A \cap E) + \bar{\beta}'(A^c \cap E)} \\ \Leftrightarrow &\frac{\beta(A \cap E) [\beta'(A \cap E) + \bar{\beta}'(A^c \cap E)]}{[\beta(A \cap E) + \bar{\beta}(A^c \cap E)] [\beta'(A \cap E) + \bar{\beta}'(A^c \cap E)]} \\ \geq &\frac{\beta'(A \cap E) [\beta(A \cap E) + \bar{\beta}(A^c \cap E)]}{[\beta(A \cap E) + \bar{\beta}(A^c \cap E)] [\beta'(A \cap E) + \bar{\beta}'(A^c \cap E)]} \\ \Leftrightarrow &\beta(A \cap E) \bar{\beta}'(A^c \cap E) \geq \beta'(A \cap E) \bar{\beta}(A^c \cap E) \\ \Leftrightarrow &\frac{\beta(A \cap E)}{1 - \beta(A \cup E^c)} \geq \frac{\beta'(A \cap E)}{1 - \beta'(A \cup E^c)} \end{aligned}$$

Hence if we find a belief function $\beta'' \in C_{\ll}(\nu)$ for which

$$\frac{\beta''(A \cap E)}{1 - \beta''(A \cup E^c)} = \max_{\beta \in C_{\ll}(\nu)} \frac{\beta(A \cap E)}{1 - \beta(A \cup E^c)}$$

then we may conclude that

$$\max_{\beta \in C_{\ll}(\nu)} \frac{\beta(A \cap E)}{\beta(A \cap E) + \bar{\beta}(A^c \cap E)} = \frac{\beta''(A \cap E)}{\beta''(A \cap E) + \bar{\beta}''(A^c \cap E)}$$

As a putative candidate for such an argmax, consider the following set function:

$$\beta_{E,A}(B) := \begin{cases} 0 & \text{if } (A \cap E) \not\subseteq B \\ \nu(A \cap E) & \text{if } (A \cap E) \subseteq B \text{ and } (A \cup E^c) \not\subseteq B \\ \nu(A \cup E^c) & \text{if } (A \cup E^c) \subseteq B \text{ and } B \neq S \\ 1 & \text{if } B = S \end{cases}$$

Notice that $\beta_{E,A}$ is a belief function. Furthermore, by construction, $\beta_{E,A}(B) \leq \nu(B)$ for all $B \in \mathcal{E}$, and so $\beta_{E,A} \in C_{\ll}(\nu)$. Thus

$$\max_{\beta \in C_{\ll}(\nu)} \frac{\beta(A \cap E)}{1 - \beta(A \cup E^c)} \geq \frac{\beta_{E,A}(A \cap E)}{1 - \beta_{E,A}(A \cup E^c)} = \frac{\nu(A \cap E)}{1 - \nu(A \cup E^c)}.$$

But since for every $\beta \in C_{\ll}(\nu)$, $\beta(A \cap E) \leq \nu(A \cap E)$ and $\beta(A \cup E^c) \leq \nu(A \cup E^c)$, it follows that

$$\max_{\beta \in C_{\ll}(\nu)} \frac{\beta(A \cap E)}{1 - \beta(A \cup E^c)} \leq \frac{\nu(A \cap E)}{1 - \nu(A \cup E^c)}.$$

Hence we have

$$\max_{\beta \in C_{\ll}(\nu)} \frac{\beta(A \cap E)}{1 - \beta(A \cup E^c)} = \frac{\nu(A \cap E)}{1 - \nu(A \cup E^c)}$$

and therefore

$$\max_{\beta \in C_{\ll}(\nu)} \frac{\beta(A \cap E)}{\beta(A \cap E) + \bar{\beta}(A^c \cap E)} = \frac{\beta_{E,A}(A \cap E)}{\beta_{E,A}(A \cap E) + \bar{\beta}_{E,A}(A^c \cap E)} = \frac{\nu(A \cap E)}{\nu(A \cap E) + \bar{\nu}(A^c \cap E)}.$$

■

As a corollary, if the event on which we are conditioning has a non-zero capacity, then we can show Denneberg's conditional expectation of any real valued function may be expressed as the Choquet integral wrt to the corresponding conditional capacity.

Formally we have: if $\nu(E) > 0$, then

$$E_{\nu}[X|E] = \int_E X d\nu_E$$

where $\nu_E(\cdot)$ is a capacity defined by

$$\nu_E(A) := E_\nu[\mathbf{1}_A|E].$$

To show this, it is first convenient to associate with a real-valued function $X : S \rightarrow \mathbb{R}$ with finite range, the coarsest (finite) partition over S , of the form $\{A_1, \dots, A_n\}$ to which X is measurable and *ordered*. That is, for any pair of states $s, t \in S$, if both s and t are in some $A \in \{A_1, \dots, A_n\}$ then $X(s) = X(t)$, otherwise $X(s) \neq X(t)$. Furthermore for any $s \in A_i$ and $t \in A_j$, $i < j$ implies $X(s) > X(t)$. For each $i = 1, \dots, n$, if we let x_i be outcome resulting if a state in A_i obtains, then X may be expressed in the form $X = [x_1 \text{ on } A_1; \dots; x_n \text{ on } A_n]$, where $x_1 > \dots > x_n$, for all $i \neq j$. Define $A^i := \bigcup_{j=1}^i A_j$. We have

$$\begin{aligned} X &= \begin{bmatrix} x_1 & \text{on } A_1 \\ x_2 & \text{on } A_2 \\ x_3 & \text{on } A_3 \\ \vdots & \vdots \\ x_{n-1} & \text{on } A_{n-1} \\ x_n & \text{on } A_n \end{bmatrix} = (x_1 - x_2) \begin{bmatrix} 1 & \text{on } A_1 \\ 0 & \text{on } A_2 \\ 0 & \text{on } A_3 \\ \vdots & \vdots \\ 0 & \text{on } A_{n-1} \\ 0 & \text{on } A_n \end{bmatrix} + (x_2 - x_3) \begin{bmatrix} 1 & \text{on } A_1 \\ 1 & \text{on } A_2 \\ 0 & \text{on } A_3 \\ \vdots & \vdots \\ 0 & \text{on } A_{n-1} \\ 0 & \text{on } A_n \end{bmatrix} \\ &\quad + \dots + (x_{n-1} - x_n) \begin{bmatrix} 1 & \text{on } A_1 \\ 1 & \text{on } A_2 \\ 1 & \text{on } A_3 \\ \vdots & \vdots \\ 1 & \text{on } A_{n-1} \\ 0 & \text{on } A_n \end{bmatrix} + x_n \begin{bmatrix} 1 & \text{on } A_1 \\ 1 & \text{on } A_2 \\ 1 & \text{on } A_3 \\ \vdots & \vdots \\ 1 & \text{on } A_{n-1} \\ 1 & \text{on } A_n \end{bmatrix} = \sum_{i=1}^{n-1} (x_i - x_{i+1}) \mathbf{1}_{A^i} + x_n. \end{aligned}$$

Thus by the additivity of the Choquet integral for sums of comonotonic functions it follows that for any $s \in E$,

$$\begin{aligned} E_\nu [X|E] &= \sum_{i=1}^{n-1} (x_i - x_{i+1}) E_\nu [\mathbf{1}_{A^i}|E] + x_n \\ &= \sum_{i=1}^{n-1} (x_i - x_{i+1}) \nu_E(A^i) + x_n = \int_E X d\nu_E. \end{aligned}$$

6 Comparing Three Updating Rules

In this section we compare the Full Bayesian Updating (FBU) rule with two other updating rules for capacities, namely the optimistic updating rule and the Dempster-Shafer rule. We argue that the FBU is superior since the Dempster-Shafer rule is unduly biased towards ambiguity-aversion, while the Optimistic rule is unduly biased towards ambiguity preference. The following two conditions on the relation between conditional and unconditional preferences have been shown by Gilboa & Schmeidler (1993) to lead to the Pessimistic (or Dempster-Shafer) and Optimistic Updating rules.

Axiom 4 (*Pessimistic Certainty Equivalent Consistency*) For any unconditionally non-null event $E \notin \mathcal{N}$ any outcome x in X , and any act f in \mathcal{F} , $f \sim_E x$ implies $f_E M \sim x_E M$.

Axiom 5 (*Optimistic Certainty Equivalent Consistency*) For any unconditionally non-null event $E \notin \mathcal{N}$ any outcome x in X , and any act f in \mathcal{F} , $f \sim_E x$ implies $f_E 0 \sim x_E 0$.

Definition 3 *The pessimistic or Dempster-Shafer updating rule is given by*

$$\nu_E(A) = \frac{\bar{\nu}(E) - \bar{\nu}(A^c \cap E)}{\bar{\nu}(E)}$$

The optimistic updating rule is given by

$$\nu_E(A) = \frac{\nu(A \cap E)}{\nu(E)}$$

To compare these two alternative updating rules with the FBU rule derived above, first consider the CEU of a ‘bet on A ’ $x_A y$ (where $x \succ y$).

$$V(x_A y) = \nu(A) u(x) + \bar{\nu}(A^c) u(y).$$

The CEU of that bet conditional on E having obtained is

$$V_E(x_A y) = \nu_E(A) u(x) + \bar{\nu}_E(A^c) u(y).$$

For the Optimistic Updating Rule this becomes

$$V_E^O(x_A y) = \frac{\nu(A \cap E)}{\nu(E)} u(x) + \frac{[\nu(E) - \nu(A \cap E)]}{\nu(E)} u(y).$$

For the Dempster-Shafer Updating Rule this becomes

$$V_E^{DS}(x_A y) = \frac{[\bar{\nu}(E) - \bar{\nu}(A^c \cap E)]}{\bar{\nu}(E)} u(x) + \frac{\bar{\nu}(A^c \cap E)}{\bar{\nu}(E)} u(y).$$

And for the Full Bayesian Updating Rule this becomes

$$V_E^{GB}(x_A y) = \frac{\nu(A \cap E)}{\nu(A \cap E) + \bar{\nu}(A^c \cap E)} u(x) + \frac{\bar{\nu}(A^c \cap E)}{\nu(A \cap E) + \bar{\nu}(A^c \cap E)} u(y).$$

The Optimistic Updating Rule seems to be updating the decision weight on the good outcome a la Bayes, with the decision weight on the bad outcome determined as the residual. The Dempster-Shafer Rule seems to be updating the decision weight on the bad outcome a la Bayes with the decision weight on the good outcome determined as the residual. The Full Bayesian Updating rule, on the other hand updates both the capacity determining the decision weight on the good outcome and the conjugate capacity determining the weight on

the bad outcome in a ‘balanced’ or ‘symmetric’ way. So aesthetically, it is also the most appealing rule!

Aesthetics apart, a more compelling argument in favor of the Full Bayesian Updating rule can be seen when the capacity that is to be updated exhibits both optimistic and pessimistic attitudes toward uncertainty. A particular simple and parsimoniously parameterized capacity that exhibits such behavior is the neo-additive capacity introduced by Chateauneuf, Grant and Eichberger (2004).

Neo-additive capacities may be viewed as a convex combination of an additive capacity and a special capacity that only distinguishes between whether an event is impossible, possible or certain. Since the Choquet integral of an act with respect to this special capacity corresponds to the Hurwicz criterion for decision making under uncertainty, we refer to this non-additive capacity as a Hurwicz capacity.

We begin by considering a partition of the set of events into the following three subsets; the set of ‘null’ events, the set of ‘universal’ events and the set of ‘essential’ events, denoted \mathcal{N} , \mathcal{U} and \mathcal{E}^* , respectively. Consistent with our usage above, a set is ‘null’ if ‘loosely speaking’ it is impossible for it to occur. Formally, we assume that this collection of events satisfies the following properties: (i) $\emptyset \in \mathcal{N}$, (ii) if $A \in \mathcal{N}$, then $B \in \mathcal{N}$, for all $B \subset A$ and (iii) if $A, B \in \mathcal{N}$ then $A \cup B \in \mathcal{N}$. A ‘universal’ set is one that is viewed as being certain to occur. Formally, it is the set of events obtained by taking the complements of each member of the set of null events, that is, $\mathcal{U} = \{E \in \mathcal{E} : S \setminus E \in \mathcal{N}\}$. Notice that since $\emptyset \in \mathcal{N}$, it follows from the definition of the set of universal events that $S \in \mathcal{U}$. Furthermore, if $A \in \mathcal{U}$, then from property (ii) for \mathcal{N} , it follows that if $A \subset B$ then $B \in \mathcal{U}$. And from property (iii) for \mathcal{N} , it follows that if $A, B \in \mathcal{U}$ then $A \cap B \in \mathcal{U}$. Finally, every other set is ‘essential’ in the sense that is neither impossible nor certain, that is, $\mathcal{E}^* = \mathcal{E} \setminus (\mathcal{N} \cup \mathcal{U})$.

Definition 4 *Fix a set of null events $\mathcal{N} \subset \mathcal{E}$. A capacity $\nu : \mathcal{E} \rightarrow [0, 1]$ is congruent with \mathcal{N} if $\nu(E) = 0$ and $\nu(E^c) = 1$, for all $E \in \mathcal{N}$. Furthermore the capacity is exactly congruent if $\nu(E) > 0$, for all $E \notin \mathcal{N}$*

For a given set of null events, the Hurwicz capacity is a minimally discriminating capacity that is exactly congruent with the set of null events. That is, it assigns only one of three possible values to an event depending on whether the event is ‘impossible’ (that is, null), ‘possible’ (that is, essential) or ‘certain’ (that is, universal).

Definition 5 Fix the set of null events $\mathcal{N} \subset \mathcal{E}$ and fix $\alpha \in [0, 1]$. The Hurwicz capacity exactly congruent with \mathcal{N} and with an α degree of optimism is defined to be

$$\mu_{\alpha}^{\mathcal{N}}(E) = \begin{cases} 0 & \text{if } E \in \mathcal{N} \\ \alpha & \text{if } E \notin \mathcal{N} \text{ and } S \setminus E \notin \mathcal{N} \\ 1 & \text{if } S \setminus E \in \mathcal{N} \end{cases}$$

Formally, we define a neo-additive capacity as a convex combination of a Hurwicz capacity and a congruent additive capacity.

Definition 6 For a given set of null events $\mathcal{N} \subset \mathcal{E}$, a finitely additive probability distribution π on (S, \mathcal{E}) that is congruent with \mathcal{N} and a pair of numbers $(\delta, \alpha) \in [0, 1]^2$, a neo-additive capacity $\nu(\cdot | \mathcal{N}, \pi, \delta, \alpha)$ is defined as

$$\nu(E | \mathcal{N}, \pi, \delta, \alpha) := (1 - \delta) \pi(E) + \delta \mu_{\alpha}^{\mathcal{N}}(E).$$

The weight $(1 - \delta)$ on the probability measure may be interpreted as the decision maker’s degree of confidence in his additive beliefs. The remaining weight δ may in turn be viewed as the lack of confidence in the additive beliefs, and depending on the relative degree of optimism (i.e. α) a fraction of that ‘residual’ weight is assigned to the best outcome that occurs on a non-null event with the remainder assigned to the worst outcome that occurs on a non-null event. Thus, the Choquet integral of a simple function f with respect to a neo-additive capacity, denoted as $\int u \circ f \, d\nu(\cdot | \mathcal{N}, \pi, \delta, \alpha)$, is equal to:

$$(1 - \delta) E_{\pi}[f] + \delta (\alpha \cdot \max \{x : f^{-1}(x) \notin \mathcal{N}\} + (1 - \alpha) \cdot \min \{y : f^{-1}(y) \notin \mathcal{N}\})$$

So, let us now consider a family of conditional neo-additive preferences.

Definition 7 (Neo-additive Preferences) The set of conditional CEU preference relations $\langle \succsim_E \rangle_{E \in \mathcal{E}}$ is said to constitute a collection of neo-additive preferences, if for each \succsim_E ,

$$\nu_E(A) = \begin{cases} 0 & \text{if } A \in \mathcal{N}_E \\ \delta_E \alpha_E + (1 - \delta_E) \pi_E(A) & \text{if } A \notin \mathcal{N}_E \text{ and } A^c \notin \mathcal{N}_E \\ 1 & \text{if } A^c \in \mathcal{N}_E \end{cases}$$

where $\delta_E, \alpha_E \in [0, 1]$.

As a corollary to our main representation (Theorem 1) we have the following for neo-additive preferences.

Corollary 1 For a set of consequentialist neo-additive CEU preferences Consequentialism, State Independence and Conditional Certainty Equivalent Consistency imply

$$\nu_E(A) = \begin{cases} 0 & \text{if } A \in \mathcal{N}_E \\ \delta_E \alpha + (1 - \delta_E) \pi_E(A) & \text{if } A \notin \mathcal{N}_E \text{ and } A^c \notin \mathcal{N}_E \\ 1 & \text{if } A^c \in \mathcal{N}_E \end{cases}$$

$$\text{where } \delta_E = \frac{\delta}{(1 - \delta) \pi(E) + \delta}$$

$$\text{and } \pi_E(A) = \begin{cases} \pi(A \cap E) / \pi(E) & \text{if } \pi(E) > 0 \\ 0 & \text{if } \pi(E) = 0 \end{cases}.$$

Proof. From Theorem 1 and the definition of a neo-additive capacity, we have for any pair $A, E \in \mathcal{E}$, $\nu_E(A) = 0$, if $A \in \mathcal{N}_E$ and $\nu_E(A) = 1$, if $A^c \in \mathcal{N}_E$. So consider the case where $A, A^c \notin \mathcal{N}_E$. Applying Theorem 1 we have

$$\begin{aligned} \nu_E(A) &= \frac{\nu(A \cap E)}{1 - \nu(E^c \cup A) + \nu(A \cap E)} \\ &= \frac{\delta \alpha + (1 - \delta) \pi(A \cap E)}{1 - [\delta \alpha + (1 - \delta) \pi(E^c \cup A)] + \delta \alpha + (1 - \delta) \pi(A \cap E)} \\ &= \frac{\delta \alpha + (1 - \delta) \pi(A \cap E)}{(1 - \delta) \pi(E) + \delta} \\ &= \frac{\delta}{(1 - \delta) \pi(E) + \delta} \times \alpha + \frac{(1 - \delta) \pi(E)}{(1 - \delta) \pi(E) + \delta} \times \frac{\pi(A \cap E)}{\pi(E)} \end{aligned}$$

■

What we find particularly appealing about the Full Bayesian update of the neo-additive capacity is that for each conditioning event E and the associated conditional neo-additive preference relation \succsim_E , the relative degree of optimism parameter α is unchanged, but the ‘lack of confidence’ parameter δ_E is related to the ex ante likelihood $\pi(E)$ of the conditioning event E . The less likely it was for the conditioning event to arise, the less confidence the individual attaches to the additive component of the updated neo-additive capacity.

Compare this to the updated capacities obtained by applying the Pessimistic (Dempster-Shafer) updating rule and the Optimistic updating rule. Letting ν_E^{DS} (respectively, ν_E^O) denote the updated capacity conditional on E obtaining that corresponds to the Dempster-Shafer (respectively, Optimistic) updating rule, straightforward application of the respective rule and algebraic manipulation yields:

$$\begin{aligned}\nu_E^{DS}(A) &= [1 - \delta_{E,\alpha}^{DS}] \pi_E(A), \text{ where } \delta_{E,\alpha}^{DS} = \frac{\delta(1 - \alpha)}{(1 - \delta)\pi(E) + \delta(1 - \alpha)} \\ \nu_E^O(A) &= [1 - \delta_{E,\alpha}^O] \pi_E(A) + \delta_{E,\alpha}^O, \text{ where } \delta_{E,\alpha}^O = \frac{\delta\alpha}{(1 - \delta)\pi(E) + \delta\alpha}.\end{aligned}$$

That is, with both of these rules, the degree of optimism of the updated capacity differs dramatically from the degree of optimism of the original unconditional capacity. No matter what event $E \subset S$ we condition upon, with the Dempster-Shafer (pessimistic) updating rule, $\alpha_E^{DS} = 0$ and with the Optimistic updating rule, $\alpha_E^O = 1$.

7 One-time versus Sequential Resolution of Simple Bets: ‘Good News’ versus ‘Bad News’ versus ‘No News’

In this section we apply our analysis to consider an individual’s attitude towards different information structures or signals. Up to this point, the only objects of choice we have considered have been acts that involve a single stage or one-time resolution of uncertainty. By its definition, however, an information structure or signal, necessarily entails two stages of resolution of uncertainty. In the first stage, the agent receives a signal leading her to update her beliefs over states and preferences over acts. Then in the second stage she learns

which state of the world has obtained and hence what outcome she receives. We shall model an information structure or signal as a (finite) partition of the state space. Thus the objects of choice shall be act/partition pairs of the form $(f, \{E_1, \dots, E_n\})$. The individual will learn in the first stage which element of the partition $\{E_1, \dots, E_n\}$ has obtained, and in the second stage will receive the outcome $f(s)$, where s is the state of the world that obtains in the second stage.

How might we be able to deduce the preferences of the decision maker over the set of such act/partition pairs that entail *sequential* resolution of uncertainty, from her family of unconditional and conditional (static) preferences over acts? One possibility is to apply the unconditional preference relation to the ‘reduction’ of the two-stage uncertainty, that is, to the act that results from the two-stage resolution of uncertainty. In the example above, this is of course the original act f . An alternative approach is to assume that the agent’s preferences over two-stage resolutions of uncertainty exhibit a recursive structure. In particular, one could assume that if each second-stage act is replaced by its conditional certainty equivalent, then the individual is indifferent between the resulting one-stage act and the original two-stage act.

To state these two notions formally, let $V(f)$ (respectively, $V_E(f)$) be a function that represents an agent’s unconditional (respectively, conditional on knowing that E has obtained) preferences over acts and let $V^2(f \mid \{E_1, \dots, E_n\})$ be a function that represents her preferences over act/partition pairs.

Definition 8 *The function $V^2(\cdot \mid \cdot)$ generates a preference relation over act/partition pairs that satisfies reduction if for all pairs of acts f and g , and all pairs of partitions $\{E_1, \dots, E_n\}$ and $\{F_1, \dots, F_m\}$:*

$$V^2(f \mid \{E_1, \dots, E_n\}) \geq V^2(g \mid \{F_1, \dots, F_m\}) \Leftrightarrow V(f) \geq V(g).$$

Definition 9 *The function $V^2(\cdot \mid \cdot)$ generates a preference relation over act/partition pairs that is recursive if for all pairs of acts f and g , all pairs of partitions $\{E_1, \dots, E_n\}$ and*

$\{F_1, \dots, F_m\}$, and all collections of outcomes $\langle x_1, \dots, x_n \rangle$ and $\langle y_1, \dots, y_m \rangle$:

$$V_{E_i}(x_i) = V_{E_i}(f) \text{ for all } i = 1, \dots, n \text{ and } V_{F_j}(y_j) = V_{F_j}(g) \text{ for all } j = 1, \dots, m$$

$$\text{implies } V^2(f \mid \{E_1, \dots, E_n\}) \geq V^2(g \mid \{F_1, \dots, F_m\})$$

$$\Leftrightarrow V\left((x_1)_{E_1}(x_2)_{E_2} \dots (x_{n-1})_{E_{n-1}}(x_n)\right) \geq V\left((y_1)_{F_1}(y_2)_{F_2} \dots (y_{m-1})_{F_{m-1}}(y_m)\right).$$

Recursivity may be viewed as an independence requirement for conditional preferences. Deducing a preference representation over act/partition pairs with recursivity uses properties of the conditional preferences, while deducing these preferences via reduction does not require knowledge of the conditional preferences. Hence, we view reduction as not being particularly interesting in the context of a model which takes conditional and unconditional preferences as primitives. In particular, reduction entails an equivalence in terms of preference for the individual between the one-time versus sequential resolution of uncertainty. Recursivity on the other hand potentially allows for a distinction (in terms of the ex ante preference) to be drawn between different information structures.

As an illustration of the structure of recursive preferences associated with a specific family of static preferences, consider a family of capacities $\langle \nu_E \rangle_{E \in \mathcal{E}}$ and a utility index u , with $u(M) = 1$ and $u(0) = 0$ that represent a collection of CEU preferences $\langle \succsim_E \rangle_{E \in (\mathcal{E})}$ that satisfies Consequentialism, State Independence and Conditional Certainty Equivalent Consistency.

For the act f , the Choquet expected utility of the one-time resolution of this act is simply $V(f) = \int u \circ f \, d\nu$. Let us now compare this to a situation in which the individual will first learn which element of the finite partition $\{E_1, \dots, E_n\}$ of S has obtained and then in the second stage will learn which state has obtained. Suppose that $\nu(E_i) > 0$ for each i , and that the events in the partition have been ordered in a way that agrees with the ordering of the conditional certainty equivalents of the act f . That is,

$$V_{E_i}(f) = \int u \circ f \, d\nu_{E_i} \geq \int u \circ f \, d\nu_{E_{i+1}} = V_{E_{i+1}}(f), \text{ for all } i = 1, \dots, n-1.$$

Thus the (recursive) Choquet expected utility of the two-stage resolution of uncertainty described above is given by

$$V^2(f|\{E_1^f, \dots, E_n^f\}) = \nu(E_1^f) \left(\int u \circ f \, d\nu_{E_1} \right) + \sum_{i=2}^n \left[\nu(\cup_{j=1}^i E_j^f) - \nu(\cup_{j=1}^{i-1} E_j^f) \right] \left(\int u \circ f \, d\nu_{E_i} \right).$$

Recall from Theorem 1 it follows that the individual's updated capacities conform to the Generalized Bayesian Updating rule and so we have

$$\nu_{E_i}(A) = \frac{\nu(A \cap E_i)}{\nu(A \cap E_i) + \bar{\nu}(A^c \cap E_i)}.$$

In addition, if $\nu(\cdot)$ is additive (that is, it is a probability measure) it follows that $\nu(\cdot) \equiv \bar{\nu}(\cdot)$ and $\nu(A \cap E_i) + \bar{\nu}(A^c \cap E_i) = \nu(E_i)$. So we have

$$\begin{aligned} & V^2(f|\{E_1, \dots, E_n\}) \\ &= \nu(E_1^f) \frac{\left(\int_{E_1} u \circ f \, d\nu_{E_1} \right)}{\nu(E_1^f)} + \sum_{i=2}^n \left[\nu(\cup_{j=1}^i E_j^f) - \nu(\cup_{j=1}^{i-1} E_j^f) \right] \frac{\left(\int_{E_i} u \circ f \, d\nu_{E_i} \right)}{\nu(E_i^f)} \\ &= \sum_{i=1}^n \int_{E_i} u \circ f \, d\nu_{E_i} = V(f). \end{aligned}$$

This is of course the well known result that a standard subjective expected utility maximizer is indifferent between the one-time or sequential resolution of the uncertainty. But how does this change if the individual exhibits optimistic and/or pessimistic attitudes towards ambiguity?

For concreteness let us consider the case of an entrepreneur establishing a 'firm' that will undertake a single project. The project will either payoff M or payoff zero. Let A be the event in which the project (and hence, firm) pays off M . Thus owning the firm is like holding the simple bet on the event A that corresponds to the act $M_A 0$. Let $\{E_1, E_2, E_3, E_4\}$ be a four-element partition in which $A = E_1 \cup E_2$.

Suppose at stage 1, there will be available to whomever manages the firm, a signal or information structure that corresponds to the partition $\{E_1, E_2 \cup E_3, E_4\}$. That is, if the manager learns in stage 1 that E_1 (respectively, E_4) has obtained then the manager knows

in stage 1 that the project will succeed (respectively, that it will fail). If the manager learns $E_2 \cup E_3$ has obtained then the result of the project will not be known until the end of second stage. Let us take the message space the manager has available in stage 1 to consist of three elements $\{G, B, N\}$ where G means “good news – the project will be successful”, B means “bad news – the project will fail” and N is the “null message” when the manager does not report anything about the success or failure of the project.

We shall restrict message (pure) strategies of the manager, $a : \{E_1, E_2 \cup E_3, E_4\} \rightarrow \{G, B, N\}$ to be non-deceitful. That is, we shall not allow $a(E_1) = B$, and we shall not allow $a(E_4) = G$. Moreover we shall also insist $a(E_2 \cup E_3) = N$ (since there is a chance that if $a(E_2 \cup E_3)$ was either G or B , that the manager would be found out to have been “deceitful” at the end of the second state should a state from the “wrong” event had obtained leading to success of the project when the manager said B or failure of the project when the manager had said G).

If we assume, however, that the only thing the public can verify ex post is whether the project has succeeded or failed (i.e. whether the outcome is M or 0) then we do allow for non-revealing strategies such as those that have $a(E_1) = N$ and/or $a(E_4) = N$.

With these restrictions, there are only four admissible (i.e. non-deceitful) message strategies.

1. Non-revealing Strategy $a_{NR}(\cdot)$, where $a_{NR}(E_1) = a_{NR}(E_2 \cup E_3) = a_{NR}(E_4) = N$.
2. Fully-Revealing Strategy $a_{FR}(\cdot)$, where $a_{FR}(E_1) = G$, $a_{FR}(E_2 \cup E_3) = N$, $a_{FR}(E_4) = B$.
3. Revealing Only Good News Strategy $a_{GN}(\cdot)$, where $a_{GN}(E_1) = G$, $a_{GN}(E_2 \cup E_3) = a_{GN}(E_4) = N$.
4. Revealing only Bad News Strategy $a_{BN}(\cdot)$, where $a_{BN}(E_1) = a_{BN}(E_2 \cup E_3) = N$, $a_{BN}(E_4) = B$.

We shall not worry about “incentive compatibility” issues that might arise concerning whether a manager after receipt of the message in stage 1 would actually want to follow

through with the message strategy adopted at the beginning. Indeed for expositional ease, let us imagine the founder of the firm is able to design the information revealing policy at the start, and suppose that once it has been decided there is a technology (say, external auditors or board of directors) which ensures compliance with this message strategy.

It is immediate to see that the four message strategies lead respectively to the following act/partition pairs:

1. Non-Revealing Strategy corresponds to $(M_A 0 \mid \{S\})$
2. Fully-Revealing Strategy corresponds to $(M_A 0 \mid \{E_1, E_2 \cup E_3, E_4\})$.
3. Revealing Only Good News Strategy corresponds to $(M_A 0 \mid \{E_1, E_2 \cup E_3 \cup E_4\})$.
4. Revealing Only Bad News Strategy corresponds to $(M_A 0 \mid \{E_1 \cup E_2 \cup E_3, E_4\})$.

The problem for the entrepreneur is to choose which message strategy to adopt for the firm. Suppose her objective is to maximize the sale price of the firm at the beginning. We assume that the entrepreneur sells the firm at time zero in a competitive market at a price equal to the certainty equivalent of the act/partition pair to which it corresponds. Let p_r denote the sale price, if the entrepreneur adopts the message strategy a_r for the firm, where $r \in \{NR, FR, GN, BN\}$. That is, if the entrepreneur sets up the firm with message strategy a_{NR} , then its sale price will be $u^{-1}(V^2(M_A 0 \mid \{S\}))$, et cetera.

Suppose the “public” is a neo-additive CEU maximizer who satisfies the axioms of theorem 1 and also its preferences over act/partition pairs satisfy Recursivity. Let $V(\cdot)$ be a function that represents its unconditional preferences over acts, with associated utility function $u(\cdot)$ and neo-additive capacity characterized by the probability measure $\pi(\cdot)$, and the parameters $\alpha, \delta \in (0, 1)$. Let $V_E(\cdot)$ be a function that represents its conditional preferences over acts (when it knows that the event E has obtained); and let $V^2(\cdot \mid \cdot)$ be a function that represents its preferences over act/partition pairs.

Furthermore assume $\min\{\pi(E_1), \pi(E_2), \pi(E_3), \pi(E_4)\} > 0$, that is, the “additive” part of the neo-additive capacity assigns non-zero probability to each element of the partition

$\{E_1, E_2, E_3, E_4\}$.

In the appendix we prove the following set of inequalities holds:

1. $p_{GN} > p_{NR} > p_{BN}$
2. $p_{GN} > p_{FR} > p_{BN}$
3. $p_{FR} > p_{NR} \Leftrightarrow \alpha > \pi_2 / (\pi_2 + \pi_3)$.

The reason the ‘reveal only good news’ signal is ranked top and the ‘reveal only bad news’ signal is ranked bottom is that neo-additive capacities overweight ‘extreme’ events. With the ‘reveal only good news’ strategy, the worst event at the end of stage 1 is the continuation value of the act when the event E_1 does not obtain, which is better than failure for sure, the worst eventuality of the non-revealing strategy in stage 2. Conversely, with the ‘reveal only bad news’ strategy, the best event at the end of stage 1 is the continuation value of the act when the event E_4 does not obtain, which is worse than success for sure, the best eventuality of the non-revealing strategy in stage 2. Similar reasoning leads to the inequalities in 2.

For the final inequality, to understand what determines whether the fully-revealing strategy is better than the non-revealing strategy, notice that the updated neo-additive preferences are closer to a weighting $(\alpha, 1 - \alpha)$ of success and failure, the more unlikely (according to the additive measure of the neo-additive capacity) was the conditioning event. So the ‘updated’ continuation preferences will rank the bet more favorably than the ‘unconditional’ preferences when α , the degree of optimism is greater than $\pi_2 / (\pi_2 + \pi_3)$ the conditional probability of success derived from the additive belief (conditional on the signal not revealing the outcome in stage 1.)

8 Conclusion

It is one of the attractive features of expected utility theory that Bayesian updating provides a natural method for considering information which becomes sequentially available. In contrast, for Choquet expected utility (CEU) theories of decision making, there are many

“sensible” updating rules in the literature, all of which share the property that the updating rule converges to Bayesian updating as capacities become additive. Which updating rule one chooses in an application will most likely depend on the intended application as much as on a priori criteria from decision theory and statistics.

In this paper we find that two axioms connecting conditional and unconditional CEU preferences characterize the Full Bayesian Updating rule (FBU). As we show by an example which generalizes the Ellsberg paradox, new information may either generate ambiguity or remove it altogether. Hence, a reasonable consistency condition should allow decision makers to maintain their attitudes towards ambiguity in the face of information that changes the ambiguity of an act. The two axioms, *conditional certainty equivalence consistency* and *consequentialism*, yield the FBU rule for capacities and achieve this desideratum.

Denneberg (2002) suggests an alternative approach to arrive at a reasonable updating rule for capacities. For a probability distribution the conditional expectation of an act is well-defined, if the conditioning event has positive probability. Assigning a value for the conditional expectation also in cases of events which have probability of zero, one can define the conditional expected value of an act with respect to a capacity as the largest expected value over all probability distributions dominated by the belief function with the smallest expected value among all belief functions dominating the capacity. This rule for finding a conditional expectation for a capacity is well-defined and, when applied to the indicator function, yields also the FBU rule for updating capacities.

Finally, applying the FBU rule to neo-additive capacities, a class of capacities with constant attitude towards ambiguity, we show that the attitude towards ambiguity remains unaffected by new information. In economic applications this is an important feature which does not hold for other updating rules.

We conclude the paper with a section on the value of information. For neo-additive capacities, which have the same attitude towards ambiguity for any conditioning event, one obtains the intuitive result that good news is preferred to no news which in turn is preferred to bad news. This preference may explain the fact why good news is more likely to be

released than bad news, an observation often made in practice.

The results of this paper strengthen the case for the FBU rule of updating capacities, in particular in economic applications.

A Proof of Theorem 1

(1) \Rightarrow (2). Fix a pair of events $E, B \in \mathcal{E}$, for which $\nu(B \cap E) > 0$ and $\bar{\nu}(B^c \cap E) > 0$.

Step 1. We shall show that u_E is an affine transformation of u . From State Independence it immediately follows that u_E is a monotonic transformation of u . Now, since X is a connected topological space, both u and u_E are continuous and u_E is a monotonic transformation of u , it is sufficient to show that for any x, y, z , if $u_E(x) > u_E(y)$ and $u_E(z) = [u_E(x) + u_E(y)]/2$ then $u(z) = [u(x) + u(y)]/2$.

Fix $x \succ y$. By State Independence, $u_E(x) > u_E(y)$.

Lemma 1 $\nu_E(B) \in (0, 1)$.

Proof. Suppose to the contrary that $\nu_E(B) = 0$, that is, $M_B 0 \sim_E 0$. Since by Consequentialism, $E^c \in \mathcal{N}_E$, $M_{B \cap E} 0 \sim_E 0$. And by Conditional Certainty Equivalent Consistency we also have $M_{B \cap E} 0 \sim 0$. That is, $\nu(B \cap E) = 0$, which contradicts the hypothesis that $\nu(B \cap E) > 0$. So suppose instead that $\nu_E(B) = 1$, that is, $M_B 0 \sim_E M$. By set monotonicity, it follows $M_{B \cup E^c} 0 \sim_E M$ and conditional certainty equivalent consistency entails $M_{B \cup E^c} 0 \sim M$. That is, $\nu(B \cup E^c) = 1$. But this is a contradiction since we have by hypothesis $\bar{\nu}(B^c \cap E) = 1 - \nu(B \cup E^c) > 0$. \square

Since $\nu_E(B) \in (0, 1)$, X is a connected topological space and u_E is continuous, outcomes

with the following properties exist.

$$z : u_E(z) = \frac{1}{2}u_E(x) + \frac{1}{2}u_E(y) \quad (\text{A.2})$$

$$z' : u_E(z') = \nu_E(B)u_E(x) + [1 - \nu_E(B)]u_E(z) \quad (\text{A.3})$$

$$y' : u_E(y') = \nu_E(B)u_E(z) + [1 - \nu_E(B)]u_E(y) \quad (\text{A.4})$$

$$z'' : u_E(z'') = \nu_E(B)u_E(x) + [1 - \nu_E(B)]u_E(y) \quad (\text{A.5})$$

Equations (2), (3), (4) and (5) together imply

$$x_{B \cap E} z_{B^c \cap E} z' \sim_E z' \quad (\text{A.6})$$

$$z_{B \cap E} y_{B^c \cap E} y' \sim_E y' \quad (\text{A.7})$$

$$x_{B \cap E} y_{B^c \cap E} z'' \sim_E z'' \quad (\text{A.8})$$

$$x_{B \cap E} y_{B^c \cap E} z'' \sim_E z'_{B \cap E} y'_{B^c \cap E} z'' \quad (\text{A.9})$$

Each of the indifferences (6), (7) and (8) follow directly from equations (3), (4) and (5), respectively. To see that the last indifference also follows, notice that the conditional (on E obtaining) Choquet Expected utility of the act $z'_{B \cap E} y'_{B^c \cap E} z''$ may be expressed

$$\begin{aligned} & \nu_E(B)u_E(z') + [1 - \nu_E(B)]u_E(y') \\ = & \nu_E(B)[\nu_E(B)u_E(x) + [1 - \nu_E(B)]u_E(z)] + [1 - \nu_E(B)][\nu_E(B)u_E(z) + [1 - \nu_E(B)]u_E(y)] \\ = & [\nu_E(B)]^2 u_E(x) + 2\nu_E(B)[1 - \nu_E(B)] \left[\frac{1}{2}u_E(x) + \frac{1}{2}u_E(y) \right] + [1 - \nu_E(B)]^2 u_E(y) \\ = & \nu_E(B)u_E(x) + [1 - \nu_E(B)]u_E(y). \end{aligned}$$

But $\nu_E(B)u_E(x) + [1 - \nu_E(B)]u_E(y)$ is the conditional (on E obtaining) Choquet Expected utility of the act $x_{B \cap E} y_{B^c \cap E} z''$ and so the indifference (9) holds.

By applying Conditional Certainty Equivalent Consistency to (6), (7), (8) and (9) we

obtain

$$x_{B \cap E} z_{B^c \cap E} z' \sim z' \quad (\text{A.10})$$

$$z_{B \cap E} y_{B^c \cap E} y' \sim y' \quad (\text{A.11})$$

$$x_{B \cap E} y_{B^c \cap E} z'' \sim z'' \quad (\text{A.12})$$

$$x_{B \cap E} y_{B^c \cap E} z'' \sim z'_{B \cap E} y'_{B^c \cap E} z''. \quad (\text{A.13})$$

These four indifference relations imply the following four equations

$$u(z') = \nu(B \cap E) u(x) + [\nu(B \cup E^c) - \nu(B \cap E)] u(z') + [1 - \nu(B \cup E^c)] u(z) \quad (\text{A.14})$$

$$u(y') = \nu(B \cap E) u(z) + [\nu(B \cup E^c) - \nu(B \cap E)] u(y') + [1 - \nu(B \cup E^c)] u(y) \quad (\text{A.15})$$

$$u(z'') = \nu(B \cap E) u(x) + [\nu(B \cup E^c) - \nu(B \cap E)] u(z'') + [1 - \nu(B \cup E^c)] u(y) \quad (\text{A.16})$$

$$u(z'') = \nu(B \cap E) u(z') + [\nu(B \cup E^c) - \nu(B \cap E)] u(z'') + [1 - \nu(B \cup E^c)] u(y) \quad (\text{A.17})$$

Substituting from (14) for $u(z')$ and from (15) for $u(y')$ into (17), and equating (17) with (16) to eliminate $u(z'')$, we obtain

$$\begin{aligned} & \nu(B \cap E) u(x) + [1 - \nu(B \cup E^c)] u(y) \\ = & \nu(B \cap E) \left[\frac{\nu(B \cap E) u(x) + [1 - \nu(B \cup E^c)] u(z)}{\nu(B \cap E) + 1 - \nu(B \cup E^c)} \right] \\ & + [1 - \nu(B \cup E^c)] \left[\frac{\nu(B \cap E) u(z) + [1 - \nu(B \cup E^c)] u(y)}{\nu(B \cap E) + 1 - \nu(B \cup E^c)} \right]. \end{aligned}$$

Collecting terms,

$$\begin{aligned} & ([\nu(B \cap E) + 1 - \nu(B \cup E^c)] \nu(B \cap E) - [\nu(B \cap E)]^2) u(x) \\ & + ([\nu(B \cap E) + 1 - \nu(B \cup E^c)] [1 - \nu(B \cup E^c)] - [1 - \nu(B \cup E^c)]^2) u(y) \\ = & 2\nu(B \cap E) (1 - \nu(B \cup E^c)) u(z). \end{aligned}$$

Simplifying yields

$$\nu(B \cap E) \bar{\nu}(B^c \cap E) u(x) + \nu(B \cap E) \bar{\nu}(B^c \cap E) u(y) = 2\nu(B \cap E) \bar{\nu}(B^c \cap E) u(z).$$

That is,

$$u(z) = \frac{1}{2}u(x) + \frac{1}{2}u(y).$$

as required.

Step 2. We know from step 1, that u_E is a positive affine transformation of u . So normalize by setting $u_E(0) = u(0) = 0$, and $u_E(M) = u(M) = 1$. So fix $A \in \mathcal{E}$, for which $\nu(A \cap E) + \bar{\nu}(A^c \cap E) > 0$ and define z to be the outcome for which $M_{A \cap E} 0_{A^c \cap E} z \sim_E z$. That is, $u(z) = \nu_E(A \cap E)$. Since by Consequentialism $E^c \in \mathcal{N}_E$, we also have $M_A 0 \sim_E M_{A \cap E} 0_{A^c \cap E} z$, hence $\nu_E(A) = u(z)$. By Conditional Certainty Equivalent Consistency, $M_{A \cap E} 0_{A^c \cap E} z \sim z$. There are three cases to consider,

1. $z = 0$, that is, $\nu_E(A \cap E) = 0$. Thus, $M_{A \cap E} 0_{A^c \cap E} z \sim z$ implies $\nu(A \cap E) = 0$. And since $\nu(A \cap E) + \bar{\nu}(A^c \cap E) > 0$, it must be the case, $\bar{\nu}(A^c \cap E) > 0$. Thus we have

$$\nu_E(A) = \frac{\nu(A \cap E)}{\nu(A \cap E) + \bar{\nu}(A^c \cap E)} = \frac{0}{0 + \bar{\nu}(A^c \cap E)} = 0, \text{ as required.}$$

2. $z = M$, that is, $\nu_E(A \cap E) = 1$. Thus, $M_{A \cap E} 0_{A^c \cap E} z \sim z$ implies $\nu(A \cup E^c) = 1$ or equivalently, $\bar{\nu}(A^c \cap E) = 1 - \nu(A \cup E^c) = 0$. And so, $\nu(A \cap E) > 0$ and we have

$$\nu_E(A) = \frac{\nu(A \cap E)}{\nu(A \cap E) + \bar{\nu}(A^c \cap E)} = \frac{\nu(A \cap E)}{\nu(A \cap E) + 0} = 1, \text{ as required.}$$

3. $u(z) \in (0, 1)$, that is, $\nu_E(A \cap E) \in (0, 1)$. Thus, $M_{A \cap E} 0_{A^c \cap E} z \sim z$ implies $u(z) = \nu(A \cap E) + [\nu(A \cup E^c) - \nu(A \cap E)] u(z)$, and since $\nu(A \cap E) + \bar{\nu}(A^c \cap E) > 0$, then we have

$$u(z) = \frac{\nu(A \cap E)}{\nu(A \cap E) + \bar{\nu}(A^c \cap E)} = \nu_E(A), \text{ as required.}$$

B Derivation of Inequalities from Section 7

Set $\pi_i := \pi(E_i)$, for $i = 1, \dots, 4$. For the Non-Revealing Strategy,

$$\begin{aligned} u(p_{NR}) &= V^2((M_A0) \mid \{S\}) = \nu(E_1 \cup E_2) \\ &= (1 - \delta)(\pi_1 + \pi_2) + \delta\alpha. \end{aligned}$$

For the Fully-Revealing Strategy,

$$\begin{aligned} u(p_{FR}) &= V^2((M_A0) \mid \{E_1, E_2 \cup E_3, E_4\}) \\ &= \nu(E_1) + [\nu(E_1 \cup E_2 \cup E_3) - \nu(E_1)] V_{E_2 \cup E_3}(M_A0) \\ &= (1 - \delta)\pi_1 + \delta\alpha + (1 - \delta)(\pi_2 + \pi_3) \left[\frac{(1 - \delta)\pi_2 + \delta\alpha}{(1 - \delta)(\pi_2 + \pi_3) + \delta} \right]. \end{aligned}$$

For the Revealing Only Good News Strategy,

$$\begin{aligned} u(p_{GN}) &= V^2(M_A0 \mid \{E_1, E_2 \cup E_3 \cup E_4\}) \\ &= \nu(E_1) + [1 - \nu(E_1)] V_{E_2 \cup E_3 \cup E_4}(M_A0) \\ &= (1 - \delta)\pi_1 + \delta\alpha + [(1 - \delta)(\pi_2 + \pi_3 + \pi_4) + \delta(1 - \alpha)] \left[\frac{(1 - \delta)\pi_2 + \delta\alpha}{(1 - \delta)(\pi_2 + \pi_3 + \pi_4) + \delta} \right]. \end{aligned}$$

For the Revealing Only Bad news Strategy,

$$\begin{aligned} u(p_{BN}) &= V^2(M_A0 \mid \{E_1 \cup E_2 \cup E_3, E_4\}) \\ &= \nu(E_1 \cup E_2 \cup E_3) V_{E_1 \cup E_2 \cup E_3}(M_A0) \\ &= [(1 - \delta)(\pi_1 + \pi_2 + \pi_3) + \delta\alpha] \left[\frac{(1 - \delta)(\pi_1 + \pi_2) + \delta\alpha}{(1 - \delta)(\pi_1 + \pi_2 + \pi_3) + \delta} \right]. \end{aligned}$$

Since $u(\cdot)$ is increasing in p , it follows that $p_r > p_{r'}$ if and only if $u(p_r) > u(p_{r'})$.

(i) Proof of $p_{GN} > p_{NR}$.

$$\begin{aligned}
& u(p_{GN}) - u(p_{NR}) \\
&= [(1-\delta)(\pi_2 + \pi_3 + \pi_4) + \delta(1-\alpha)] \left[\frac{(1-\delta)\pi_2 + \delta\alpha}{(1-\delta)(\pi_2 + \pi_3 + \pi_4) + \delta} \right] - (1-\delta)\pi_2 \\
&= (1-\delta) \left[\frac{(\pi_2 + \pi_3 + \pi_4)(1-\delta)\pi_2 + (\pi_2 + \pi_3 + \pi_4)\delta\alpha + \delta(1-\alpha)\pi_2 + \delta^2(1-\alpha)\alpha/(1-\delta)}{(1-\delta)(\pi_2 + \pi_3 + \pi_4) + \delta} \right] \\
&\quad - (1-\delta) \left[\frac{(\pi_2 + \pi_3 + \pi_4)(1-\delta)\pi_2 + \delta\pi_2}{(1-\delta)(\pi_2 + \pi_3 + \pi_4) + \delta} \right] \\
&= \frac{(1-\delta) [(\pi_3 + \pi_4)\delta\alpha + \delta^2(1-\alpha)\alpha/(1-\delta)]}{[(1-\delta)(\pi_2 + \pi_3 + \pi_4) + \delta]} \\
&= \frac{\delta\alpha [(1-\delta)(\pi_3 + \pi_4) + \delta(1-\alpha)]}{[(1-\delta)(\pi_2 + \pi_3 + \pi_4) + \delta]} > 0.
\end{aligned}$$

(ii) Proof of $p_{NR} > p_{BN}$.

$$\begin{aligned}
& u(p_{NR}) - u(p_{BN}) \\
&= (1-\delta)(\pi_1 + \pi_2) + \delta\alpha \\
&\quad - [(1-\delta)(\pi_1 + \pi_2 + \pi_3) + \delta\alpha] \left[\frac{(1-\delta)(\pi_1 + \pi_2) + \delta\alpha}{(1-\delta)(\pi_1 + \pi_2 + \pi_3) + \delta} \right] \\
&= [(1-\delta)(\pi_1 + \pi_2) + \delta\alpha] \left[\frac{(1-\delta)(\pi_1 + \pi_2 + \pi_3) + \delta - (1-\delta)(\pi_1 + \pi_2 + \pi_3) - \delta\alpha}{(1-\delta)(\pi_1 + \pi_2 + \pi_3) + \delta} \right] \\
&= \frac{[(1-\delta)(\pi_1 + \pi_2) + \delta\alpha]\delta(1-\alpha)}{(1-\delta)(\pi_1 + \pi_2 + \pi_3) + \delta} > 0.
\end{aligned}$$

(iii) Proof of $p_{GN} > p_{FR}$.

$$\begin{aligned}
& u(p_{GN}) - u(p_{FR}) \\
&= [(1-\delta)(\pi_2 + \pi_3 + \pi_4) + \delta(1-\alpha)] \left[\frac{(1-\delta)\pi_2 + \delta\alpha}{(1-\delta)(\pi_2 + \pi_3 + \pi_4) + \delta} \right] \\
&\quad - (1-\delta)(\pi_2 + \pi_3) \left[\frac{(1-\delta)\pi_2 + \delta\alpha}{(1-\delta)(\pi_2 + \pi_3) + \delta} \right] \\
&= [(1-\delta)\pi_2 + \delta\alpha] \left[\frac{\delta(1-\alpha)}{(1-\delta)(\pi_2 + \pi_3 + \pi_4) + \delta} \right] \\
&\quad + [(1-\delta)\pi_2 + \delta\alpha](1-\delta) \left[\frac{(\pi_2 + \pi_3 + \pi_4)(1-\delta)(\pi_2 + \pi_3) + (\pi_2 + \pi_3 + \pi_4)\delta}{[(1-\delta)(\pi_2 + \pi_3 + \pi_4) + \delta][(1-\delta)(\pi_2 + \pi_3) + \delta]} \right] \\
&\quad - [(1-\delta)\pi_2 + \delta\alpha](1-\delta) \left[\frac{(\pi_2 + \pi_3)(1-\delta)(\pi_2 + \pi_3 + \pi_4) + (\pi_2 + \pi_3)\delta}{[(1-\delta)(\pi_2 + \pi_3 + \pi_4) + \delta][(1-\delta)(\pi_2 + \pi_3) + \delta]} \right]
\end{aligned}$$

$$\begin{aligned}
&= [(1-\delta)\pi_2 + \delta\alpha] \delta \left[\frac{(1-\alpha)}{(1-\delta)(\pi_2 + \pi_3 + \pi_4) + \delta} + \frac{(1-\delta)\pi_4\alpha}{[(1-\delta)(\pi_2 + \pi_3 + \pi_4) + \delta][(1-\delta)(\pi_2 + \pi_3) + \delta]} \right] \\
&> 0.
\end{aligned}$$

(iv) Proof of $p_{FR} > p_{BN}$.

$$\begin{aligned}
&u(p_{FR}) - u(p_{BN}) \\
&= (1-\delta)\pi_1 + \delta\alpha + (1-\delta)(\pi_2 + \pi_3) \left[\frac{(1-\delta)\pi_2 + \delta\alpha}{(1-\delta)(\pi_2 + \pi_3) + \delta} \right] \\
&\quad - [(1-\delta)(\pi_1 + \pi_2 + \pi_3) + \delta\alpha] \left[\frac{(1-\delta)(\pi_1 + \pi_2) + \delta\alpha}{(1-\delta)(\pi_1 + \pi_2 + \pi_3) + \delta} \right] \\
&= [(1-\delta)\pi_1 + \delta\alpha] \left[1 - \frac{(1-\delta)(\pi_1 + \pi_2) + \delta\alpha}{(1-\delta)(\pi_1 + \pi_2 + \pi_3) + \delta} \right] \\
&\quad + (1-\delta)(\pi_2 + \pi_3) \left[\frac{(1-\delta)\pi_2 + \delta\alpha}{(1-\delta)(\pi_2 + \pi_3) + \delta} - \frac{(1-\delta)(\pi_1 + \pi_2) + \delta\alpha}{(1-\delta)(\pi_1 + \pi_2 + \pi_3) + \delta} \right] \\
&= \frac{[(1-\delta)\pi_1 + \delta\alpha][(1-\delta)\pi_3 + \delta(1-\alpha)]}{(1-\delta)(\pi_1 + \pi_2 + \pi_3) + \delta} \\
&\quad + \frac{(1-\delta)(\pi_2 + \pi_3)[(1-\delta)^2\pi_2(\pi_1 + \pi_2 + \pi_3) + (1-\delta)\pi_2\delta + \delta\alpha(1-\delta)(\pi_1 + \pi_2 + \pi_3) + \delta^2\alpha]}{[(1-\delta)(\pi_2 + \pi_3) + \delta][(1-\delta)(\pi_1 + \pi_2 + \pi_3) + \delta]} \\
&\quad - \frac{(1-\delta)(\pi_2 + \pi_3)[(1-\delta)^2(\pi_1 + \pi_2)(\pi_2 + \pi_3) + (1-\delta)(\pi_1 + \pi_2)\delta + \delta\alpha(1-\delta)(\pi_2 + \pi_3) + \delta^2\alpha]}{[(1-\delta)(\pi_2 + \pi_3) + \delta][(1-\delta)(\pi_1 + \pi_2 + \pi_3) + \delta]} \\
&= \frac{[(1-\delta)\pi_1 + \delta\alpha][(1-\delta)\pi_3 + \delta(1-\alpha)]}{(1-\delta)(\pi_1 + \pi_2 + \pi_3) + \delta} \\
&\quad + \frac{(1-\delta)(\pi_2 + \pi_3)[(1-\delta)^2(-\pi_1\pi_3) + \delta(1-\delta)(-\pi_1) + \delta\alpha(1-\delta)\pi_1]}{[(1-\delta)(\pi_2 + \pi_3) + \delta][(1-\delta)(\pi_1 + \pi_2 + \pi_3) + \delta]} \\
&= \frac{[(1-\delta)\pi_1 + \delta\alpha][(1-\delta)\pi_3 + \delta(1-\alpha)]}{(1-\delta)(\pi_1 + \pi_2 + \pi_3) + \delta} - \frac{(1-\delta)^2(\pi_2 + \pi_3)\pi_1[(1-\delta)\pi_3 + \delta(1-\alpha)]}{[(1-\delta)(\pi_2 + \pi_3) + \delta][(1-\delta)(\pi_1 + \pi_2 + \pi_3) + \delta]} \\
&= \frac{[(1-\delta)\pi_3 + \delta(1-\alpha)]}{[(1-\delta)(\pi_1 + \pi_2 + \pi_3) + \delta]} \left[(1-\delta)\pi_1 + \delta\alpha - \frac{(1-\delta)^2(\pi_2 + \pi_3)\pi_1}{[(1-\delta)(\pi_2 + \pi_3) + \delta]} \right] \\
&= \frac{[(1-\delta)\pi_3 + \delta(1-\alpha)]}{[(1-\delta)(\pi_1 + \pi_2 + \pi_3) + \delta]} \left[\frac{(1-\delta)^2\pi_1(\pi_2 + \pi_3) + (1-\delta)\pi_1\delta - (1-\delta)^2(\pi_2 + \pi_3)\pi_1}{[(1-\delta)(\pi_2 + \pi_3) + \delta]} + \delta\alpha \right] \\
&= \frac{[(1-\delta)\pi_3 + \delta(1-\alpha)]\delta}{[(1-\delta)(\pi_1 + \pi_2 + \pi_3) + \delta]} \left[\frac{(1-\delta)\pi_1}{[(1-\delta)(\pi_1 + \pi_2 + \pi_3) + \delta]} + \alpha \right] > 0.
\end{aligned}$$

(v) Proof of $p_{FR} > p_{NR} \Leftrightarrow \alpha > \pi_2 / (\pi_2 + \pi_3)$.

$$\begin{aligned} u(p_{FR}) - u(p_{NR}) &= (1 - \delta) \left[(\pi_2 + \pi_3) \left[\frac{(1 - \delta) \pi_2 + \delta \alpha}{(1 - \delta) (\pi_2 + \pi_3) + \delta} \right] - \pi_2 \right] \\ &= \frac{\delta (1 - \delta) [\alpha (\pi_2 + \pi_3) - \pi_2]}{(1 - \delta) (\pi_2 + \pi_3) + \delta} > 0 \Leftrightarrow \alpha > \frac{\pi_2}{\pi_2 + \pi_3}. \end{aligned}$$

References

- Baron, J. and D. Frisch (1988): “Ambiguity and Rationality,” *Journal of Behavioral Decision Making*, **1**, 149-157.
- Denneberg, D. (1994): “Conditioning (Updating) Non-additive Measures,” *Annals of Operations Research* **52**, 21-42.
- Denneberg, D. (2002): “Conditional Expectation for Monotone Measures, the Discrete Case,” *Journal of Mathematical Economics* **37**, 105-121.
- Dow J. and S. Werlang (1992): “Uncertainty Aversion, Risk Aversion, and the Optimal Choice of Portfolio,” *Econometrica*, **60**, 197-204
- Eichberger, J. and D. Kelsey (2000): “Non-additive Beliefs and Strategic Equilibria,” *Games and Economic Behavior*, **30**, 183-215.
- Eichberger, J. and D. Kelsey (2002): “Strategic Complements, Substitutes, and Ambiguity: The Implications for Public Goods,” *Journal of Economic Theory*, **106**, 436-66.
- Eichberger, J., S. Grant and D. Kelsey (2005): “CEU Preferences and Dynamic Consistency,” *Mathematical Social Sciences*, **49**, 143-151.
- Epstein, E. and M. Schneider (2003): “Recursive Multiple-Priors,” *Journal of Economic Theory*, **113**, 1-31.
- Gilboa, I. and D. Schmeidler (1989): “Maxmin Expected Utility with a Non-unique Prior,” *Journal of Mathematical Economics* **18**, 141-153.
- Gilboa, I. and D. Schmeidler (1993): “Updating Ambiguous Beliefs,” *Journal of Economic Theory* **59**, 33-49.
- Grant, S., A. Kajii and B. Polak (2000): “Decomposable Choice,” *Journal of Economic Theory* **92**, 169-197 (doi:10.1006/jeth.2000.2644).
- Gul, F. (1991): “A Theory of Disappointment Aversion,” *Econometrica* **59**, 667-686.
- Hanany E. and P. Klibanoff (2005): “Updating Preferences with Multiple Priors,” mimeo Kellogg Graduate School of Management, Northwestern University.

- Jaffray, J.-Y. (1992): "Bayesian Updating and Belief Functions." *IEEE Transactions on Systems, Man and Cybernetics* **22**, 1144-1152.
- Keynes, J.M. (1921): *A Treatise on Probability*. Macmillan, London.
- Knight, F.H. (1921): *Risk, Uncertainty and Profit*. Houghton Mifflin, New York.
- Lapied, A. and R. Kast (2005): "Updating Choquet Valuation and Discounting Information Arrivals." mimeo CNRS, Lameta et Idep.
- Lehrer, E. (2004): "Updating Non-additive Probabilities – A Geometric Approach," mimeo School of Mathematical Sciences, Tel Aviv University.
- Mukerji, S. (1998): "Ambiguity Aversion and Incompleteness of Contractual Form," *American Economic Review*, **88**, 1207-1231.
- Mukerji, S. and J.-M. Tallon (2001): "Ambiguity Aversion and Incompleteness of Financial Markets," *Review of Economic Studies*, **68**, 883-904.
- Mukerji, S. and J.-M. Tallon (2004): "Ambiguity Aversion and the Absence of Wage Indexation," *Journal of Monetary Economics*, **51**, 653-70
- Palacios-Huerta, I. (1999): "The Aversion to the Sequential Resolution of Uncertainty," *Journal of Risk and Uncertainty* **18**, 249-269.
- Pires, C.P. (2002): "A Rule for Updating Ambiguous Beliefs," *Theory and Decision* **53**, 137-152.
- Sarin R. and P. Wakker (1998): "Dynamic Choice and Non-expected Utility," *Journal of Risk and Uncertainty* **17**, 87-119.
- Schmeidler, D. (1989): "Subjective Probability and Expected Utility without Additivity," *Econometrica* **57**, 571-587.
- Shafer, G. (1976): *A Mathematical Theory of Evidence*. Princeton University Press, Princeton New Jersey.
- Siniscalchi, M. (2004): "Dynamic Choice under Ambiguity," mimeo Department of Economics, Northwestern University (downloadable from <http://faculty.econ.northwestern.edu/faculty/siniscalchi/>)
- Skiadas, C. (1997): "Conditioning and Aggregation of Preferences," *Econometrica* **65**, 347-367.
- Walley, P. (1991): *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London.

Wakker, P. (2001): “Testing and Characterizing Properties of Nonadditive Measures through Violations of the Sure-Thing Principle,” *Econometrica* **69**, 1039-1059.

Wang, T. (2003): “Conditional Preferences and Updating,” *Journal of Economic Theory* **108**, 286-321.